

MIT Press • Open Encyclopedia of Cognitive Science

The Language of Thought Hypothesis

Nicolas Porot¹ Eric Mandelbaum²

¹Africa Institute for Research in Economics and Social Sciences, Mohammed VI Polytechnic University,

²Baruch College & The Graduate Center, City University of New York

MIT Press

Published on: Sep 10, 2025

DOI: <https://doi.org/10.21428/e2759450.79b53c0b>

License: [Creative Commons Attribution-NonCommercial 4.0 International License \(CC-BY-NC 4.0\)](#).

The language of thought hypothesis is a thesis about the structure of mental representations. It is an example of the *computational-representational theory of mind*, according to which much of cognition consists in formal computations over mental representations. What distinguishes the language of thought hypothesis from other such theories is the idea that mental representations share core features with formal languages. The language of thought hypothesis states that thinking is the transformation of mental representations in a language of thought, in which most of what we call “thinking” is in fact inferences performed over sentences in the language. The language of thought is not a natural language, such as Spanish or Haitian Creole. It is a formal language, comparable to the machine language of a computer or formal languages of logic. The language of thought is often claimed to be innate, universal, and pivotal in learning, and it is implicated in processes such as natural language acquisition. However, it is also compatible with learning and reasoning in nonlinguistic creatures as well as in mental processes seemingly unconnected to language such as perception and motor control.

History

Predecessors of the language of thought hypothesis (LoTH) can be found as early as in Plato ([Moss, 2014](#)) and Aristotle ([Panaccio, 1999](#)), continuing through the medieval period, most elaborately in the work of William of Ockam ([Panaccio, 1999](#)). The idea reappeared in the later 20th century via philosophers such as Gilbert Harman ([1973](#)) and Wilfred Sellars ([1975](#)). The LoTH as a hypothesis in cognitive science truly took shape with Jerry Fodor, who gave the first book length treatment of the topic, as its most prominent proponent ([Fodor, 1975](#)). Since then, it has been expanded and evolved with numerous proponents of the view, in philosophy ([Fodor & Pylyshyn, 1988](#); [Quilty-Dunn et al., 2023](#)), computational cognitive science ([Piantadosi et al., 2012](#); [Sablé-Meyer et al., 2022](#)), neuroscience ([Al Roumi et al., 2021](#); [Frankland & Greene, 2020](#)), and experimental psychology ([Cavanagh, 2021](#); [Dehaene et al., 2022](#)).

The 20th century resurgence of the idea that thought occurs in a mental language can be seen as part of a broader shift away from behaviorism. In the 1950s and 1960s, subdisciplines of psychology began positing mental representations in all manner of psychological explanation. There was thus a need for a better understanding of the core features of mental representations. The LoTH can be seen as a response to this need.

Early proponents of LoT, specifically Gilbert Harman and Jerry Fodor, were inspired by some of Noam Chomsky’s arguments for universal grammar. The arguments Chomsky was making for language, Fodor applied to thought: just as people would infer that language was compositional (i.e., the meaning of a linguistic expression is derived from the meaning of its parts and the way they are combined, cf. [Szabó, 2012](#)) from the fact that it was productive and systematic, so too one might infer that thought was compositional because it was productive and systematic (see the section *Chomskian-inspired LoTH arguments*).

Fodor saw the LoTH as rendering explicit what was tacit in the work of cognitive scientists at the time. To understand research that was being done in vision (e.g., feature extraction), concept learning (via hypothesis testing and confirmation), or language (acquisition, production, and understanding) [see [Visual Cognitive Neuroscience](#); [Concepts](#); [Language Acquisition](#); [Language Production](#)], one needed to appeal to LoT representations. Therefore, to even make sense of recent scientific progress, the thought went, researchers needed to posit a LoT.

Many alternatives to LoTH have targeted this conception of cognitive science. This includes other (non-LoT) variants of the computational-representational theory of mind. For example, some have argued that perceptual representations are iconic ([Block, 2023](#)), that concept learning does not require hypothesis testing and confirmation ([Carey, 2009](#)), or that computation generally is dissociable from linguistic computation ([Camp, 2007](#)). Each undermines the idea that the LoTH is the only viable account of the structure of mental representations. Contemporary defenses of the LoTH have thus tended to treat it as a “working hypothesis,” one of many possible competing views of cognitive architecture rather than as the only possible explanation.

Core concepts

Basic ingredients

All languages, including LoT, have *simple* and *complex* expressions. An atomic concept is a simple expression in a LoT. For instance, the expressions “wave” and “beach” are plausibly simple in English, as is “crash.” The expression, “The wave crashed on the beach,” is complex.

A *constituent* is a part of a complex expression. The thought, `the candidate from Trenton has a toupee`, is composed of constituents `the candidate from Trenton` and `has a toupee`.¹ The expression `has a toupee` is in turn composed of `has` and `a toupee`. Not any arbitrary part is a constituent; `Trenton has`, for example, is not a constituent of that thought. That is because there is a canonical way that sentences are broken down: `Trenton` is a constituent of `the candidate from Trenton`, and `has` is a constituent of `has a toupee`, and it is only these larger strings that are the constituents of the full expression.

Discrete constituents are typical of language-like representations, but some kinds of representation, such as icons, lack them. The complex phrase “a red dog” represents a red dog, as does a picture of a red dog. However, although the word “red” can be easily removed from the linguistic expression, the red cannot so easily be removed from the picture of the dog ([Green & Quilty-Dunn, 2021](#)). The part of the picture that instantiates redness is the same one that instantiates the dog shape.

Syntax is the set of rules governing combinations of constituents. It determines which combinations are well formed and which are not. According to the LoTH, thought is governed by syntactic rules, like language. The syntactic rules of thought explain why people can think `the wave crashed on the beach` but not `beach on crashed the wave the`. They also explain why `Louise called Abdel` means something different from `Abdel called Louise`, although they share the same constituents.

Natural languages are *compositional*. The meaning of a string in natural language is (in many cases) determined by the meanings of its parts and the way those parts are combined. According to the LoTH, thought shares this important feature with language. To illustrate, consider the thought `the wave crashed on the beach`. It is similar to, but not that same as, the thought `the wave receded on the beach`. They are similar in that they share the same syntactic structure. However, they differ in their constituents and have different meanings. Similarly, `Abdel and Louise went rock climbing` differs from both of the sentences above. It has a different syntactic structure and also has distinct constituents (`went`, `rock climbing`).

Chomskian-inspired LoTH arguments

Historically, arguments for the LoTH focused on properties of a LoT that were similar to those postulated by Chomsky for natural language.

The argument from productivity ([Fodor, 1975](#), pp. 31-32) starts from the observation that people can, in principle, think an infinite number of novel thoughts of arbitrary length despite having a finite number of concepts. This supports the existence of a mental language because the only way to build an infinite set of representations out of a finite base of representations is via recursively defined rules for combining discrete constituents.

Given that computing time and resources are bounded, the argument from productivity relies on a claim about what the mind is capable of in principle, not what any particular human actually computes. That is, the argument from productivity invokes a performance/competence distinction. The *performance* of thinking is limited by computing constraints such as the length of a person's life. But the underlying machinery, or *competence*, of thinking is not limited in these ways. One reason for supposing that people's ability to think is unbounded in principle is that rules for building thoughts look to be repeatable. If a person can think that Jo went to the store, and that Jo went to the park, they can think that Jo went to the store, then to the park, then to the store again, then to the park again, and so on, it would seem, ad infinitum. Another piece of evidence is that people can think completely novel thoughts. Consider: How many kilograms of fresh citrus are shipped to Antarctic research stations annually?

A key piece of evidence for these claims about thought stems from natural language: people can produce and understand arbitrarily long sentences and sentences they have never heard before, and both processes appear to implicate thought. Independent arguments in favor of productivity in the absence of natural language have focused on specific animal behaviors such as the caching behavior of Western scrub jays ([Gallistel & King, 2009](#)) or baboons' tracking of social dominance relations (cf. [Camp, 2009](#); [Cheney & Seyfarth, 2008](#)). Although controversial, this evidence suggests some animals can iterate their representations of relations productively. As a baboon sets up its hierarchical tree of which members and families constitute which place in the social hierarchy, members and layers of the social hierarchy can be added ad infinitum so that each new member has a space in the animal's representation of its troop's social hierarchy.

One problem for productivity arguments is that cognition is in fact bounded by time and resources, such as working memory and organism lifespan, so there is no direct evidence for productivity. Those who deny that thought has a language-like structure can reject that it is in-principle productive, citing the lack of direct evidence for infinitely complex thoughts. Some have denied that natural language is productive ([Ziff, 1974](#)) and have also raised doubts about productivity in nonhuman thought ([Dennett, 1989](#)).

The *argument from systematicity* ([Fodor & Pylyshyn, 1988](#)) is similar to the argument from productivity, but it does not rely on speculations about in-principle boundlessness and so is less susceptible to the objections concerning the use of the performance/competence distinction. It takes as its starting point the semantic relatedness of thoughts with similar syntactic structure. Consider the following pair of thoughts: Louise called Abdel , and Abdel called Louise . They appear to have the same structure and the same parts, but the parts are arranged differently. Similarly, they have almost the same meaning. The argument from systematicity claims that this parallelism, which arises for virtually any relation one can think of, is not a coincidence: thoughts have real syntactic structure, and changes in that structure cause changes in meaning.

The argument in schematic form is as follows: for any thought of the form aRb that one can think, there is another thought of the form bRa that one can also think that is semantically related to the first. On the argument from systematicity, what explains this generalization is the fact that thought, like language, has canonical parts ([Fodor, 2007](#); [Green & Quilty-Dunn, 2021](#)) that can be combined and rearranged.

Evidence for systematicity comes from patterns in natural language production and understanding, which are plausibly systematic (cf. [Johnson, 2004](#)), as well as from introspection. People have an ability to produce and understand *sentences* that show this pattern, such as “Louise called Abdel,” and “Abdel called Louise.” To explain this, researchers infer back to the kinds of thoughts people can think, which would also have to be systematic. One limitation for this argument is that it does not show systematicity of thought for prelinguistic humans or animals, since researchers lack introspective access to their thoughts, and they lack natural language. A parallel argument for systematicity to cover these cases uses learning rather than language use ([Fodor & Pylyshyn, 1988](#)): for any relation R, a creature that can learn aRb relations can also learn bRa relations. Therefore, representation of relations is systematic.

Argument from cognitive science

The LoTH is not the only computationalist account of cognition, and its ability to explain the results and data of cognitive science offers another argument for adopting it over those rival views. Since early in the history of cognitive science, researchers have proposed alternative, nonlinguistic representational formats in thought, including icons ([Greenberg, 2023](#); [Shepard & Metzler, 1971](#)), maps ([Camp, 2007](#); [Tolman, 1948](#)), vector representations ([Günther et al., 2019](#)), associations ([Mandelbaum, 2016, 2020](#)), and analog magnitudes ([Meck & Church, 1983](#)), among others. Partly for this reason, many proponents of the LoTH (including Fodor) are pluralists about kinds of mental representation, allowing that LoTs and nonlinguistic representations may coexist. This style of argument is similar in spirit to other arguments

for the LoTH but differs in important ways. Unlike other arguments, it does not require that LoT is the only possible explanation for human cognitive capacities. Rather, for each piece of evidence advanced for LoT, it relies on comparing LoT-style explanations directly to others available in the literature. Contemporary proponents have developed the argument from cognitive science by relying on discussions of large bodies of data, rather than in-principle arguments, to argue for the centrality of the LoT to cognitive science ([Quilty-Dunn et al., 2023](#); see the section *Questions, controversies, and new developments*).

Questions, controversies, and new developments

A new argument from cognitive science

Early iterations of the LoTH said little about the fundamental properties of LoT. The more recent approach identifies a cluster of language-like features that form a natural kind, one whose properties can be identified through experimental methods. Similarly, most early accounts did not specify which LoT symbols were most likely to be atomic, and which derived, or specify the most basic principles of composition. LoT was typically discussed in the singular, suggesting that it might be a unique language shared across biological cognition. Recent research has instead broadened the focus, wondering how many LoTs there might be ([Mandelbaum et al., 2022](#)). For example, if the LoTH is true of adult humans, it is almost certainly true of infants and other species, and it likely applies to more than one cognitive system within the adult mind. Each of these languages could potentially have a different evolutionary history and function(s), with different resulting computational needs and constraints. African gray parrots might use a different LoT in reasoning than a human 12 month old, and the adult human visual system might operate on yet another LoT, each with different atomic concepts, syntactic rules, or working memory constraints.

One example of this new argument from cognitive science in action comes from the domain of computational cognitive science, particularly Bayesian models of cognition. These accounts marry LoT-style sentences with a probabilistic inductive engine to explain human abilities in a wide variety of cases, including the acquisition of numerical concepts, Boolean concepts, intuitive theories, normative rules, taxonomies, and programs. Probabilistic languages of thought (PLOTs) have also been fruitful in modeling causality judgments, analogical reasoning, and mapping sentences to logical form (see [Quilty-Dunn et al., 2023](#)). Specific PLOT models of learning can also be tested against developmental data ([Carcassi & Szymanik, 2023](#)).

Properties of a LoT

Recent accounts of the LoTH have proposed six canonical properties that are characteristic of LoTs ([Quilty-Dunn et al., 2023](#)). Some LoTs may lack one or more of these properties, and some representational schemes may have just a few of these properties, making it unclear whether they are LoTs or not. However, the properties are thought to typically cluster together in a homeostatic property

cluster ([Boyd, 1990](#)): they do not always appear together, but the appearance of one increases the probability of finding the others, too. These properties allow for LoTs to be identified in experimental contexts because the features individually yield specific behavioral and neural predictions.

First, LoTs have discrete constituents (see the section *Basic ingredients*).

Second, they exhibit *role-filler independence*. A distinctive feature of LoTs is that they maintain a separation between constituents and their syntactic roles. In `Abdel drove Louise`, for example, there are roles for an agent (Abdel), an action (driving), and a patient (Louise). `Abdel` maintains its identity across changes in role, for example, if it fills the patient role in `Louise drove Abdel`.

Third, a LoT has a structure in which a predicate is applied to an argument yielding a truth-evaluable expression. For example, in the sentence “pigs fly,” the predicate “fly” is applied to the argument “pigs.” The resulting string can be evaluated as false. Predicate-argument structure is typical of languages generally. Not all well-formed formulae in a LoT are truth evaluable (e.g., imperatives like `shut the door` may be well formed but not truth apt).

Fourth, LoTs have logical operators like `not` and `if...then`, which structure deductive inferences. Logical operators have purely functional meanings: unlike “rock,” there is no set of tangible things in the world that instantiates `if...then`. Evidence for the presence of such operators is evidence for linguistic elements in thought. Other types of mental representation, such as maps ([Camp, 2007](#)) or imagistic simulation ([Leahy et al., 2022](#)), can approximate logical reasoning, such as when a mental model may simulate negation not with a logical operator like “not” but instead by just representing absence (compare “there is no rabbit in the park” to showing a picture of a rabbit-less park). Thus, special care must be taken to demonstrate that logical connectives, rather than these nonlinguistic representations, structure reasoning. Importantly, logical operators are separable from predicate-argument structure (such as in `Louise drove Abdel`); a language might allow for logical operations over representations that lack predicate-argument structure, such as maps (so that it is possible to think thoughts of the form `[map 1]` or `[map 2]`, in which `map 1` and `map 2` are cognitive maps). It might also allow for predicate-argument structure without the ability to conjoin sentences with logical operators (so that it is possible to think `it's cold` and also think `there are rabbits in the park`, but it is not possible to think `it's cold` and `there are rabbits in the park`).

Fifth, if LoTs structure thought, they should structure transitions between thoughts, too. In effect, one should expect that there are automatic, structure-sensitive transitions between states in many parts of the mind ([Quilty-Dunn & Mandelbaum, 2018](#)). To take a mundane example, one might effortlessly transition from the thought that there is no more coffee at home, and the thought that when there is no more coffee at home, one should stop at the store on the way back from work, to the thought that one should stop at the store on the way back from work. These transitions are determined by the formal properties of mental representations along with the basic inferential rules of reasoning ([Quilty-Dunn & Mandelbaum,](#)

[2018](#)) in the same way that the transitions between states in a classical computer language are. This is sometimes referred to as *inferential promiscuity*.

Sixth and finally, a LoT supports *abstract content*. People can think about many topics, from the swift (mis)carriage of justice to the wingspan of the chimney swift. Some things that people can think about, such as the concept of justice, are difficult to represent without the arbitrary relation of linguistic symbols (such as the word “justice”) to their referents. Evidence for representation of such categories is evidence for linguistic symbols in thought and thus for LoT(s). Imagistic and iconic modes of thought, traditionally the trade of empiricist theories of learning, have trouble representing abstraction. These theories often posit a learning trajectory that predicts sensorily grounded concepts to be semantically more primitive than abstract concepts and acquired before them. By contrast, since languages represent abstract concepts with symbols, they do not need to represent them using complexes of sensory representations. Abstract concepts can even be atomic. Thus, the more evidence there is for semantically primitive abstract concepts, the more one should expect LoT formats of representation.

Evidence for the LoTH

Current empirical debates around the LoTH focus on the six features identified above (see the section *Properties of a LoT*), particularly the role of logical operators and inferential promiscuity. Uncovering when logical operators are present in thought is one of the most pressing questions in LoTH and nativism debates more generally. A commonly studied operator is sentential negation (*not*). In sentential negation, an operator combines with a sentence to produce a new sentence with the opposite truth value of the first. “The Nile is the longest river on earth” (a true sentence) can be negated, becoming “The Nile is not the longest river on earth” (a false sentence). Some hold that sentential negation may be present in thought at 12 months or younger ([Cesana-Arlotti et al., 2018](#)) or may not arise until 17 to 19 months ([Feiman et al., 2022](#)). The lower bound on negation understanding in language appears to be as early as 18 months ([de Carvalho et al., 2021](#)), which serves as the upper bound for when negation is present in prelinguistic thought. Negation is a very stringent test for the development of operators, being a harder operator to process than others, such as conjunction (*and*).

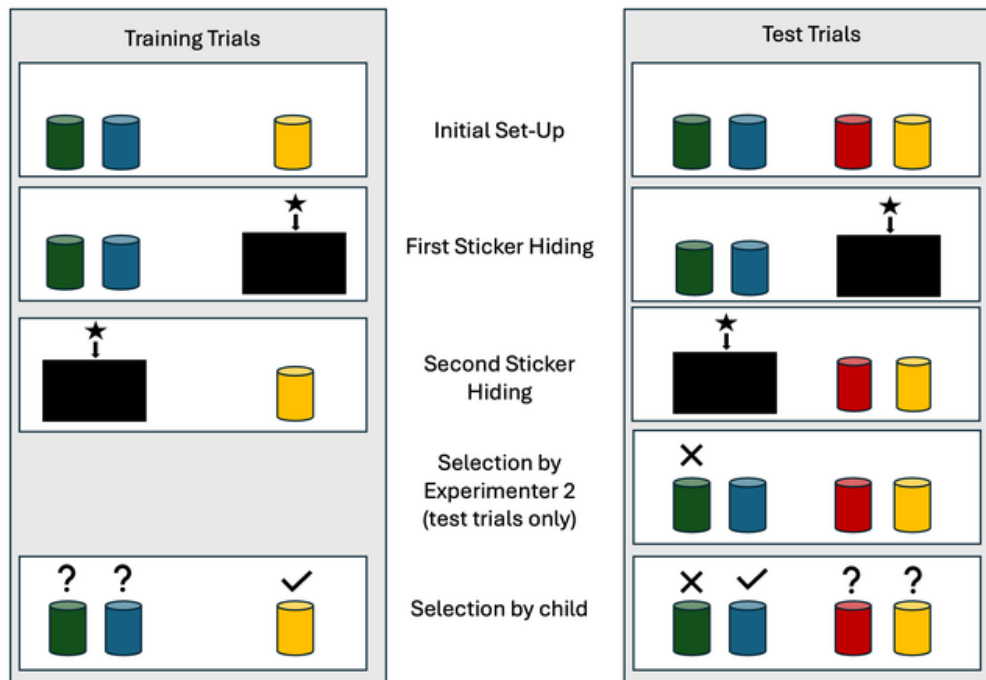


Figure 1

Schematic representation of test and training trials for a four-cup task with children, using sticker rewards. Adapted from [Mody and Carey \(2016\)](#).

Linguistic tests are not suitable for deciding when operators are present in prelinguistic human infants and animals. One area of current debate surrounds the presence of such operators in children younger than 17 months and in nonhuman animals ([Engelmann et al., 2023](#); [Leahy et al., 2022](#)). Much work focuses on *disjunctive syllogistic reasoning*, logical inferences of the form $A \text{ or } B; \text{ not } A; \text{ therefore } B$. This is used as a test for the presence of negation (*not*) and disjunction (*or*). Although there are many ways to operationalize disjunctive syllogistic reasoning (e.g., [Cesana-Arlotti & Halberda, 2024](#)), perhaps the most broadly discussed paradigm is based on Josep Call's (2004) cups task ([Mody & Carey, 2016](#); see [Figure 1](#)). In a standard version of this paradigm, participants are familiarized with a setup in which a reward may be present in one of two cups. They are then shown that one of the cups is empty. Experimenters observe whether the infants and animals successfully leverage the certainty offered by reasoning by exclusion to go for the other cup, rather than the one shown empty.

Success on the two cups tasks has been challenged as evidence of true disjunctive syllogism, as subjects may succeed via a nonlogical operator-involving strategy such as mapping unknown to unknown or avoiding empty objects ([Feiman et al., 2022](#)). In the four cups task, subjects see four cups, separated in two sets of two, with two rewards given, one in each subgroup. In this setup, subjects are shown that one cup of one pair is empty. Disjunctive syllogism would then imply to subjects exactly where one reward is, leaving the location of the reward in the other pair of cups uncertain. Whether the cups experiments show inference by disjunctive syllogism or just mental simulation (or similar nonlogical operator-involving strategies) is a topic of current research (see, e.g., [Feiman et al., 2022](#)).

Evidence for automatic inferential transitions is particularly indicative of LoT for aspects of cognition that are most remote from natural language processing and conscious control such as unconscious instantaneous reasoning. Thus, LoTs appear to be in tension with traditional dual process models of cognition that identify System 1 processing with automatic, unconscious, associative transitions in thought ([Kahneman, 2011](#)). By contrast, they are well suited to interact with inferential mechanisms, which are thought to be typical of slower, deliberative System 2 processes. Since LoT is a language, it can be used to define the formal rules of those inferences (for example, the rule *modus ponens*: if A, then B; A; therefore B).

Recent work in the dual processes literature suggests LoT processing even in System 1 processes. To take just one example, it has long been observed that beliefs can skew the ability to evaluate the quality of logical arguments ([Markovits & Nantel, 1989](#)). Humans are biased to accept invalid arguments that have conclusions they agree with and to reject good ones that have conclusions they disagree with—a typical System 1-type effect often called the *belief bias*. It appears that there is also an inverse effect of logical evaluation on belief: when asked to evaluate the truth of conclusions, participants' responses show implicit sensitivity to whether the argument was logically valid, with them having slower reaction times (and higher error rates) on problems in which logic and belief conflict. For instance, they will be, on average, slower to identify a false conclusion when it follows from a valid argument than an invalid one. This effect is robust. It persists when researchers place participants under cognitive load (to ensure System 1 processing) and even when researchers tell participants to focus on the believability of the conclusion rather than to evaluate whether the argument is logically valid ([Howarth et al., 2021](#)). This suggests that System 1 processes are at least partly logically structured, operating in a LoT.

Possible alternatives to a LoT

Connectionist networks have long been cited as an alternative to LoT architectures. Such models, also known as (artificial) neural networks, are made up of large networks of interconnected nodes and do not explicitly characterize the mind in terms of sentences in a LoT. Recent years have witnessed increasing optimism about such models with remarkable successes of deep learning, for example, in large language models, computer vision, and image- and video-generating models [see [Large Language Models](#)]. In one sense, framing things in an adversarial way is simply inaccurate, since connectionist architectures could implement a LoT ([Fodor & Pylyshyn, 1988](#)), making room for both in a single cognitive process. In another sense, however, there is a fairly deep challenge raised by such proposals. If connectionist networks can be built to mimic the performance that is used as evidence of computationalist architectures, but without implementing such architectures, why should cognitive scientists need computational-representational theory of mind at all? Why not think the mind is simply a biologically implemented connectionist network with no intervening layers of computation? Yet, if the LoTH is correct, this objection misses the point because connectionist networks lack the kind of symbolic, formal processes that are fundamental to cognition.

This debate has a rich history and is now primed for an exciting new chapter, as the power of computational cognitive science, including large language models and PLoTs, has increased dramatically. Current points of contention include (1) whether learning models could underlie the widespread prepared learning seen in humans and other animals and (2) whether some features of cognition that appear to be fundamental (such as feature processing in vision and compositionality in thought) can be approximated to a satisfying degree by learning models. These questions interact with questions of domain-general learning mechanisms (i.e., those that can be used to process any domain, such as attention, rather than being dedicated to just one domain, such as natural language). Current connectionist-inspired architectures are domain general, and often when one finds a connectionist architecture that can mimic human performance in a given domain, it is too brittle to port to other cognitive processes.

Putting these implementation questions to one side, the LoTH promises to be a productive area of research in coming years. In particular, the idea that there might be many languages of thought raises questions for theorists and experimentalists. One of these is how LoTs might interface with one another and with other formats. By way of example, it is possible that there are representations that figure in higher cognition but could never be computed over by the motor system, even in principle, because of their format. However, it is also possible that there are no such representations. Similarly, there is much to explore about the phylogenetic and ontogenetic origins of LoTs. One pressing question here is how much overlap there is between human and nonhuman LoTs. Another concerns which syntactic features of thought are most plausible for a given species or stage of human development.

Broader connections

Because the LoTH is a thesis about the underlying structure of cognitive representations, it touches virtually every subdiscipline of cognitive science. One of the promising aspects of the current state of this field is its empirical tractability. There are fruitful questions concerning the implementation of LoTs in human nervous systems ([Kazanina & Poeppel, 2023](#)). Neuroscientists can investigate what kinds of implementation models are most plausible and well supported for the LoTH ([Frankland & Greene, 2020](#)). Similarly, given advances in computational approaches to cognition, researchers can ask what kinds of artificial systems might be capable of implementing LoTs or what kinds of artificial systems might reliably mimic compositionality ([McGrath et al., 2023](#)).

The fact that LoTs can differ from one another raises the possibility that different systems in the mind might operate over different LoTs. Visual perception might operate in one or more LoTs, for example, while logical operations in reasoning operate over another [see [Visual Cognitive Neuroscience](#); [Causal Reasoning](#)]. There is even more room for exploration using the LoTH framework in development and comparative cognition [see [Cognitive Development](#); [Animal Cognition](#)]. The syntactic rules and atomic concepts in the minds of infants and nonhuman animals are largely unknown, and testing between hypotheses in such cases will require an extensive amount of careful experimentation. The LoTH allows

for careful comparison of different logical operators and quantifiers. For instance, one can ask about the role of generics (statements such as “Birds are cute” as opposed to the specific “That bird is cute”) in LoTs, such as why they are processed more easily than existential and universally quantified statements. Perhaps it is because they are a default mode of thought (Rhodes et al., 2025).

There are questions regarding the phylogeny and ontogeny of LoTs. When, how, and how many separate times they emerged will be an area in which comparative research will be indispensable; developmental researchers can address whether and how LoTs change in grammar and expressive power through infancy and childhood and how such changes might shape development (Kibbe, 2023). A firmer grasp of the predicates of the infant mind can also elucidate which concepts are likely to be atomic and thus potentially unlearned, offering new purchase on classic debates about nativism, empiricism, and the origins of abstract content in thought.

Acknowledgments

Helpful comments were received from Xander Macswan, Jessica Moss, Griffin Pion, Jake Quilty-Dunn, Marjorie Rhodes, and Elliot Schwartz, who are all hereby thanked for their contributions.

Further reading

- Dehaene, S., Al Roumi, F., Lakretz, Y., Planton, S., & Sablé-Meyer, M. (2022). Symbols and mental programs: A hypothesis about human singularity. *Trends in Cognitive Sciences*, 26(9), 751-766. <https://doi.org/10.1016/j.tics.2022.06.010>
- Fodor, J. A. (1975). *The language of thought*. MIT.
- Kazanina, N., & Poeppel, D. (2023). The neural ingredients for a language of thought are available. *Trends in Cognitive Sciences*, 27(11), 996-1007. <https://doi.org/10.1016/j.tics.2023.07.012>
- Quilty-Dunn, J., Porot, N., & Mandelbaum, E. (2022). The best game in town: The re-emergence of the language of thought hypothesis across the cognitive sciences. *Behavioral and Brain Sciences*, 46, e261. <https://doi.org/10.1017/S0140525X22002849>

Footnotes

1. We use this font change, with gray background, to indicate expressions in a LoT. Double quotes signify an expression in natural language. ↩

References

- Al Roumi, F., Marti, S., Wang, L., Amalric, M., & Dehaene, S. (2021). Mental compression of spatial sequences in human working memory using numerical and geometrical primitives. *Neuron*, 109(16), 2627-2639.e4. <https://doi.org/10.1016/j.neuron.2021.06.009>

- Block, N. (2023). *The border between seeing and thinking*. Oxford University Press.

↩

- Boyd, R. (1990). Realism, approximate truth, and philosophical method. In C. W. Savage (Ed.), *Scientific theories* (pp. 355–391). University of Minnesota Press.

↩

- Call, J. (2004). Inferences about the location of food in the great apes (*Pan paniscus*, *Pan troglodytes*, *Gorilla gorilla*, and *Pongo pygmaeus*). *Journal of Comparative Psychology*, 118(2), 232–241.
<https://doi.org/10.1037/0735-7036.118.2.232>

↩

- Camp, E. (2007). Thinking with maps. *Philosophical Perspectives*, 21(1), 145–182.
<https://doi.org/10.1111/j.1520-8583.2007.00124.x>

↩

- Camp, E. (2009). A language of baboon thought. In R. W. Lurz (Ed.), *The philosophy of animal minds* (pp. 108–127). Cambridge University Press.

↩

- Carcassi, F., & Szymanik, J. (2023). The Boolean language of thought is recoverable from learning data. *Cognition*, 239, 105541. <https://doi.org/10.1016/j.cognition.2023.105541>

↩

- Carey, S. (2009). *The origin of concepts*. Oxford University Press.

↩

- Cavanagh, P. (2021). The language of vision. *Perception*, 50(3), 195–215.
<https://doi.org/10.1177/0301006621991491>

↩

- Cesana-Arlotti, N., & Halberda, J. (2024). A continuity in logical development: Domain-general disjunctive inference by toddlers. *Open Mind*, 8, 809–825. https://doi.org/10.1162/opmi_a_00148

↩

- Cesana-Arlotti, N., Martín, A., Téglás, E., Vorobyova, L., Cetnarski, R., & Bonatti, L. L. (2018). Precursors of logical reasoning in preverbal human infants. *Science*, 359(6381), 1263–1266.
<https://doi.org/10.1126/science.aao3539>

↩

- Cheney, D.L. & Seyfarth, R.M. (2008). *Baboon Metaphysics: The Evolution of a Social Mind*. Chicago University Press.

↩

- de Carvalho, A., Crimon, C., Barrault, A., Trueswell, J., & Cristophe, A. (2021). “Look. It is not a bamoule!”: 18- and 24-month-olds can use negative sentences to constrain their interpretation of novel word meanings. *Developmental Science*, 24(4), e13085. <https://doi.org/10.1111/desc.13085>

↩

- Dehaene, S., Al Roumi, F., Lakretz, Y., Planton, S., & Sablé-Meyer, M. (2022). Symbols and mental programs: A hypothesis about human singularity. *Trends in Cognitive Sciences*, 26(9), 751-766. <https://doi.org/110.1016/j.tics.2022.06.010>

↩

- Dennett, D.C. (1989). Mother nature versus the walking encyclopedia: A Western drama. In W. Ramsey, S. P. Stich, & D. E. Rumelhart (Eds.), *Philosophy and connectionist theory* (pp. 21-30). Erlbaum.

↩

- Engelmann, J. M., Haux, L. M., Völter, C., Schleihauf, H., Call, J., Rakoczy, H., & Herrmann, E. (2023). Do chimpanzees reason logically? *Child Development*, 94, 1102–1116. <https://doi.org/10.1111/cdev.13861>

↩

- Feiman, R., Mody, S., & Carey, S. (2022). The development of reasoning by exclusion in infancy. *Cognitive psychology*, 135, 101473. <https://doi.org/10.1016/j.cogpsych.2022.101473>

↩

- Fodor, J. A. (1975). *The language of thought*. Harvard University Press.

↩

- Fodor, J. A. (2007). The revenge of the given. In B. P. McLaughlin & J. Cohen (Eds.), *Contemporary debates in philosophy of mind* (pp. 105-116). Wiley-Blackwell.

↩

- Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28(1-2), 3–71. [https://doi.org/10.1016/0010-0277\(88\)90031-5](https://doi.org/10.1016/0010-0277(88)90031-5)

↩

- Frankland, S. M., & Greene, J. D. (2020). Concepts and compositionality: in search of the brain's language of thought. *Annual review of psychology*, 71(1), 273-303.

↩

- Gallistel, C. R., & King, A. P. (2009). *Memory and the computational brain: Why cognitive science will transform neuroscience*. Wiley-Blackwell. <https://doi.org/10.1002/9781444310498>

↩

- Green, E. J., & Quilty-Dunn, J. (2021). What is an object file? *British Journal for the Philosophy of Science*, 72(3), 665-699.

↩

- Greenberg, G. (2023). The iconic-symbolic spectrum. *Philosophical Review*, 132(4), 579-627. <https://doi.org/10.1215/00318108-10697558>

↩

- Günther, F., Rinaldi, L., & Marelli, M. (2019). Vector-space models of semantic representation from a cognitive perspective: A discussion of common misconceptions. *Perspectives on Psychological Science*, 14(6), 1006-1033.

↩

- Harman, G. (1973). *Thought*. Princeton University Press.

↩

- Howarth, S., Handley, S., & Polito, V. (2021). Uncontrolled logic: Intuitive sensitivity to logical structure in random responding. *Thinking & Reasoning*, 28(1), 1-36. <https://doi.org/10.1080/13546783.2021.1934119>

↩

- Johnson, K. (2004). On the systematicity of language and thought. *Journal of Philosophy*, 101(3), 111-139. <https://doi.org/10.5840/jphil2004101321>

↩

- Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

↩

- Kazanina, N., & Poeppel, D. (2023). The neural ingredients for a language of thought are available. *Trends in Cognitive Sciences*, 27(11), 996-1007. <https://doi.org/10.1016/j.tics.2023.07.012>

↩

- Kibbe, M. M. (2023). The language-of-thought as a working hypothesis for developmental cognitive science. *Behavioral and Brain Sciences*, 46, e280. <https://doi.org/10.1017/S0140525X23002030>

↩

- Leahy, B., Huemer, M., Steele, M., Alderete, S., & Carey, S. (2022). Minimal representations of possibility at age 3. *Proceedings of the National Academy of Sciences*, 119(52), e2207499119.

<https://doi.org/10.1073/pnas.2207499119>

↩

- Mandelbaum, E. (2016). Attitude, inference, association: On the propositional structure of implicit bias. *Noûs*, 50(3), 629-658. <https://doi.org/10.1111/nous.12089>

↩

- Mandelbaum, E. (2020). Associationist theories of thought. In E. N. Zalta & U. Nodelman (Eds.), *Stanford encyclopedia of philosophy*. <https://plato.stanford.edu/archives/fall2025/entries/associationist-thought/>

↩

- Mandelbaum, E., Dunham, Y., Feiman, R., Firestone, C., Green, E. J., Harris, D., Kibbe, M. M., Kurdi, B., Mylopoulos, M., Shepherd, J., Wellwood, A., Porot, N., & Quilty-Dunn, J. (2022). Problems and mysteries of the many languages of thought. *Cognitive Science*, 46(12), e13225. <https://doi.org/10.1111/cogs.13225>

↩

- Markovits, H., & Nantel, G. (1989). The belief-bias effect in the production and evaluation of logical conclusions. *Memory & Cognition*, 17(1), 11-17. <https://doi.org/10.3758/bf03199552>

↩

- McGrath, S. W., Russin, J., Pavlick, E., & Feiman, R. (2023). Properties of LoTs: The footprints or the bear itself? *Behavioral and Brain Sciences*, 46, e284. <https://doi.org/10.1017/S0140525X23001863>

↩

- Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, 9(3), 320-324.

↩

- Mody, S., & Carey, S. (2016). The emergence of reasoning by the disjunctive syllogism in early childhood. *Cognition*, 154, 40-48. <https://doi.org/10.1016/j.cognition.2016.05.012>

↩

- Moss, J. (2014). X—Plato's Appearance-Assent Account of Belief. *Proceedings of the Aristotelian Society*, 114(2), 213-238. <https://doi.org/10.1111/j.1467-9264.2014.00371.x>

↩

- Panaccio, C. (1999). *Mental language: From Plato to William of Okham*. Fordham University Press.

↩

- Piantadosi, S. T., Tenenbaum, J. B., & Goodman, N. D. (2012). Bootstrapping in a language of thought: A formal model of numerical concept learning. *Cognition*, 123(2), 199-217.

<https://doi.org/10.1016/j.cognition.2011.11.005>

↩

- Quilty-Dunn, J., Porot, N., & Mandelbaum, E. (2023). The best game in town: The re-emergence of the language of thought hypothesis across the cognitive sciences. *Behavioral and Brain Sciences*, 46, e261. <https://doi.org/10.1017/S0140525X22002849>

↩

- Quilty-Dunn, J., & Mandelbaum, E. (2018). Inferential transitions. *Australasian Journal of Philosophy*, 96(3), 532-547. <https://doi.org/10.1080/00048402.2017.1358754>

↩

- Rhodes, M., Gelman, S., & Leslie, S. (2025). How generic language shapes the development of social thought. *Trends in Cognitive Science*, 29(2), 122-132. <https://doi.org/10.1016/j.tics.2024.09.012>

↩

- Sablé-Meyer, M., Ellis, K., Tenenbaum, J., & Dehaene, S. (2022). A language of thought for the mental representation of geometric shapes. *Cognitive Psychology*, 139, 101527. <https://doi.org/10.1016/j.cogpsych.2022.101527>

↩

- Sellars, W. (1975). The structure of knowledge. In H.-N. Castañeda (Ed.), *Action, knowledge, and reality*. Bobbs-Merrill.

↩

- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171(3972), 701–703. <https://doi.org/10.1126/science.171.3972.701>

↩

- Szabó, Z. G. (2012). The case for compositionality. In W. Hinzen, E. Machery, & M. Werning (Eds.), *The Oxford handbook of compositionality*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199541072.013.0003>

↩

- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55(4), 189–208. <https://doi.org/10.1037/h0061626>

↩

- Ziff, P. (1974). The number of English sentences. *Foundations of Language* 11(4), 519–532.

↩